

Ref #	Hits	Search Query	DBs	Default Operator	Plurals	Time Stamp
L1	4	("6,778,979") or ("6,732,090") or ("6,611,825") or ("6,233,575")).PN.	US-PGPUB; USPAT	OR	OFF	2005/02/06 18:42

Welcome to IEEE Xplore®

- Home
- What Can I Access?
- Log-out

Tables of Contents

- Journals & Magazines
- Conference Proceedings
- Standards

Search

- By Author
- Basic
- Advanced
- CrossRef

Member Services

- Join IEEE
- Establish IEEE Web Account
- Access the IEEE Member Digital Library

IEEE Enterprise

- Access the IEEE Enterprise File Cabinet

 Print Format


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
Search: The ACM Digital Library The Guide


[Feedback](#) [Report a problem](#) [Satisfaction survey](#)
Terms used [natural language](#) and [data mining](#) and [goal](#) and [Bayesian](#)

Found 7,770 of 150,138

Sort results
by

[Save results to a Binder](#)
[Try an Advanced Search](#)
Display
results

[Search Tips](#)
[Try this search in The ACM Guide](#)
 [Open results in a new window](#)

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale

1 [Multi Relational Data Mining \(MRDM\): Probabilistic logic learning](#)

Luc De Raedt, Kristian Kersting

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1Full text available: [pdf\(1.98 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

The past few years have witnessed an significant interest in probabilistic logic learning, i.e. in research lying at the intersection of probabilistic reasoning, logical representations, and machine learning. A rich variety of different formalisms and learning techniques have been developed. This paper provides an introductory survey and overview of the state-of-the-art in probabilistic logic learning through the identification of a number of important probabilistic, logical and learning concept ...

Keywords: data mining, inductive logic programming, machine learning, multi-relational data mining, probabilistic reasoning, uncertainty

2 [Survey articles: Data mining for hypertext: a tutorial survey](#)

Soumen Chakrabarti

January 2000 **ACM SIGKDD Explorations Newsletter**, Volume 1 Issue 2Full text available: [pdf\(1.19 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

With over 800 million pages covering most areas of human endeavor, the World-wide Web is a fertile ground for data mining research to make a difference to the effectiveness of information search. Today, Web surfers access the Web through two dominant interfaces: clicking on hyperlinks and searching via keyword queries. This process is often tentative and unsatisfactory. Better support is needed for expressing one's information need and dealing with a search result in more structured ways than av ...

3 [Computing curricula 2001](#)

September 2001 **Journal on Educational Resources in Computing (JERIC)**Full text available: [pdf\(613.63 KB\)](#) [html\(2.78 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

4

[The KDD process for extracting useful knowledge from volumes of data](#)

Usama Fayyad, Gregory Piatetsky-Shapiro, Padhraic Smyth
November 1996 **Communications of the ACM**, Volume 39 Issue 11

Full text available:  [pdf\(523.49 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)



5 Industrial/government track: Clinical and financial outcomes analysis with existing hospital patient records

R. Bharat Rao, Sathyakama Sandilya, Radu Stefan Niculescu, Colin Germond, Harsha Rao
August 2003 **Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(188.40 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Existing patient records are a valuable resource for automated outcomes analysis and knowledge discovery. However, key clinical data in these records is typically recorded in unstructured form as free text and images, and most structured clinical information is poorly organized. Time-consuming interpretation and analysis is required to convert these records into structured clinical data. Thus, only a tiny fraction of this resource is utilized. We present REMIND, a Bayesian Framework for Reliable ...

Keywords: Bayes Nets, HMMs, data mining, temporal reasoning



6 KM-1 (knowledge management): clustering I: Goal-oriented methods and meta methods for document classification and their parameter tuning

Stefan Siersdorfer, Sergej Sizov, Gerhard Weikum
November 2004 **Proceedings of the Thirteenth ACM conference on Information and knowledge management**

Full text available:  [pdf\(228.34 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Automatic text classification methods come with various calibration parameters such as thresholds for probabilities in Bayesian classifiers or for hyperplane distances in SVM classifiers. In a given application context these parameters should be set so as to meet the relative importance of various result quality metrics such as precision versus recall. In this paper we consider classifiers that can accept a document for a topic, reject it, or abstain. We aim to meet the application's goals in ...

Keywords: meta classification, restrictive classification



7 Panel and workshop reports from KDD-2003: Multirelational data mining 2003: workshop report

Saso Dzeroski, Luc De Raedt, Stefan Wrobel
December 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 2

Full text available:  [pdf\(45.28 KB\)](#) Additional Information: [full citation](#), [abstract](#)

In this report, we briefly review the second International Workshop on Multi-Relational Data Mining (MRDM-03), which was organized by the authors and held in Washington, D.C. on August 27th, 2003 as part of the workshop program of the ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-03). the goal of the workshop was to bring together researchers and practitioners of Data Mining and interested in methods and applications of finding patterns in expressive langu ...

Keywords: multi-relation learning and data mining

8 Textual data mining of service center call records

Pang-Ning Tan, Hannah Blau, Steve Harp, Robert Goldman

August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(178.04 KB\)](#) Additional Information: [full citation](#), [references](#), [index terms](#)

9 Multi Relational Data Mining (MRDM): Biological applications of multi-relational data mining

David Page, Mark Craven

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1

Full text available:  [pdf\(1.12 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Biological databases contain a wide variety of data types, often with rich relational structure. Consequently multi-relational data mining techniques frequently are applied to biological data. This paper presents several applications of multi-relational data mining to biological data, taking care to cover a broad range of multi-relational data mining techniques.

10 Strategic directions in artificial intelligence

Jon Doyle, Thomas Dean

December 1996 **ACM Computing Surveys (CSUR)**, Volume 28 Issue 4

Full text available:  [pdf\(243.02 KB\)](#) Additional Information: [full citation](#), [references](#), [index terms](#)

11 Special issue on the fusion of domain knowledge with data for decision support: Fusion of domain knowledge with data for structural learning in object oriented domains

Helge Langseth, Thomas D. Nielsen

December 2003 **The Journal of Machine Learning Research**, Volume 4

Full text available:  [pdf\(227.18 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

When constructing a Bayesian network, it can be advantageous to employ structural learning algorithms to combine knowledge captured in databases with prior information provided by domain experts. Unfortunately, conventional learning algorithms do not easily incorporate prior information, if this information is too vague to be encoded as properties that are local to families of variables. For instance, conventional algorithms do not exploit prior information about repetitive structures, which are ...

12 Bioinformatics—an introduction for computer scientists

Jacques Cohen

June 2004 **ACM Computing Surveys (CSUR)**, Volume 36 Issue 2

Full text available:  [pdf\(261.56 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

The article aims to introduce computer scientists to the new field of bioinformatics. This area has arisen from the needs of biologists to utilize and help interpret the vast amounts of data that are constantly being gathered in genomic research---and its more recent counterparts, proteomics and functional genomics. The ultimate goal of bioinformatics is to develop *in silico* models that will complement *in vitro* and *in vivo* biological experiments. The article provides a bird's eye view of the ...

Keywords: DNA, Molecular cell biology, RNA and protein structure, alignments, cell simulation and modeling, computer, dynamic programming, hidden-Markov-models, microarray, parsing biological sequences, phylogenetic trees

13 Special issue on on inductive logic programming: ilp: a short look back and a longer look forward

David Page, Ashwin Srinivasan

December 2003 **The Journal of Machine Learning Research**, Volume 4

Full text available: [pdf\(103.21 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Inductive logic programming (ILP) is built on a foundation laid by research in machine learning and computational logic. Armed with this strong foundation, ILP has been applied to important and interesting problems in the life sciences, engineering and the arts. This paper begins by briefly reviewing some example applications, in order to illustrate the benefits of ILP. In turn, the applications have brought into focus the need for more research into specific topics. We enumerate and elaborate f ...

14 Scalable association-based text classification

Dimitris Meretakis, Dimitris Fragoudis, Hongjun Lu, Spiros Likothanassis

November 2000 **Proceedings of the ninth international conference on Information and knowledge management**

Full text available: [pdf\(149.74 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: machine learning and IR, statistical/probabilistic models, text categorization, text data mining

15 Technical Papers: Applying natural language processing (NLP) based metadata extraction to automatically acquire user preferences

Woojin Paik, Sibel Yilmazel, Eric Brown, Maryjane Poulin, Stephane Dubon, Christophe Amice
October 2001 **Proceedings of the international conference on Knowledge capture**

Full text available: [pdf\(210.42 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

This paper describes a metadata extraction technique based on natural language processing (NLP) which extracts personalized information from email communications between financial analysts and their clients. Personalized means connecting users with content in a personally meaningful way to create, grow, and retain online relationships. Personalization often results in the creation of user profiles that store individuals' preferences regarding goods or services offered by various e-commerce merch ...

Keywords: metadata extraction, natural language processing, user preference elicitation

16 Discovering models of software processes from event-based data

Jonathan E. Cook, Alexander L. Wolf

July 1998 **ACM Transactions on Software Engineering and Methodology (TOSEM)**,
Volume 7 Issue 3

Full text available: [pdf\(369.76 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Many software process methods and tools presuppose the existence of a formal model of a process. Unfortunately, developing a formal model for an on-going, complex process can be difficult, costly, and error prone. This presents a practical barrier to the adoption of process technologies, which would be lowered by automated assistance in creating formal models. To this end, we have developed a data analysis technique that we term process discovery. Under this technique, data ...

Keywords: Balboa, process discovery, software process, tools

17 Empirical bayes screening for multi-item associations

William DuMouchel, Daryl Pregibon

August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**Full text available:  [pdf\(931.67 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

This paper considers the framework of the so-called "market basket problem", in which a database of transactions is mined for the occurrence of unusually frequent item sets. In our case, "unusually frequent" involves estimates of the frequency of each item set divided by a baseline frequency computed as if items occurred independently. The focus is on obtaining reliable estimates of this measure of interestingness for all item sets, even item sets with relatively low frequencies. For example, in ...

Keywords: Association rules, Data Mining, Knowledge Discovery, Statistical Models, empirical Bayes methods, gamma-Poisson model, market basket problem, shrinkage estimation

18 Research track papers: Probabilistic author-topic models for information discovery

Mark Steyvers, Padhraic Smyth, Michal Rosen-Zvi, Thomas Griffiths

August 2004. **Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining**Full text available:  [pdf\(323.72 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

We propose a new unsupervised learning technique for extracting information from large text collections. We model documents as if they were generated by a two-stage stochastic process. Each author is represented by a probability distribution over topics, and each topic is represented as a probability distribution over words for that topic. The words in a multi-author paper are assumed to be the result of a mixture of each authors' topic mixture. The topic-word and author-topic distributions are ...

Keywords: Gibbs sampling, text modeling, unsupervised learning

19 Probabilistic query models for transaction data

Dmitry Pavlov, Padhraic Smyth

August 2001 **Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining**Full text available:  [pdf\(958.33 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

We investigate the application of Bayesian networks, Markov random fields, and mixture models to the problem of query answering for transaction data sets. We formulate two versions of the querying problem: the query selectivity estimation (i.e., finding exact counts for tuples in a data set) and the query generalization problem (i.e., computing the probability that a tuple will occur in new data). We show that frequent itemsets are useful for reducing the original data to a compressed representa ...

20 An evaluation of statistical spam filtering techniques

Le Zhang, Jingbo Zhu, Tianshun Yao

December 2004 **ACM Transactions on Asian Language Information Processing (TALIP)**,

Volume 3 Issue 4

Full text available:  [pdf\(343.64 KB\)](#)Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

This paper evaluates five supervised learning methods in the context of statistical spam filtering. We study the impact of different feature pruning methods and feature set sizes on each learner's performance using cost-sensitive measures. It is observed that the significance of feature selection varies greatly from classifier to classifier. In particular, we found support vector machine, AdaBoost, and maximum entropy model are top performers in this evaluation, sharing similar characteristics: ...

Keywords: Spam filtering, text categorization

Results 1 - 20 of 200

Result page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)


[Subscribe \(Full Service\)](#) [Register \(Limited Service, Free\)](#) [Login](#)
Search: The ACM Digital Library The Guide

[Feedback](#) [Report a problem](#) [Satisfaction survey](#)
Terms used [natural language](#) and [data mining](#) and [goal](#) and [Bayesian](#)

Found 7,770 of 150,138

Sort results by

[Save results to a Binder](#)
Display results

[Search Tips](#)
 [Open results in a new window](#)
[Try an Advanced Search](#)
[Try this search in The ACM Guide](#)
Results 21 - 40 of 200 Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

Best 200 shown

Relevance scale

21 [Bioinformatics \(BIO\): BioMap: toward the development of a knowledge base of biomedical literature](#)

 Kamal Kumar, Mathew J. Palakal, Snehasis Mukhopadhyay, Mathew J. Stephens, Huian Li
 March 2004 [Proceedings of the 2004 ACM symposium on Applied computing](#)

 Full text available: [pdf\(212.77 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

Biological literature databases continue to grow rapidly with vital information that is important for conducting sound biomedical research. As data and information space continue to grow exponentially, the need for rapidly surveying the published literature, synthesizing, and discovering the embedded "knowledge" is becoming critical to allow the researchers to conduct "informed" work, avoid repetition, and generate new hypotheses. Knowledge, in this case, is defined as one-to-many and many-to-many ...

Keywords: bioinformatics, data mining, databases, machine learning, text mining

22 [Using information scent to model user information needs and actions and the Web](#)

Ed H. Chi, Peter Pirolli, Kim Chen, James Pitkow

 March 2001 [Proceedings of the SIGCHI conference on Human factors in computing systems](#)

 Full text available: [pdf\(278.29 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

On the Web, users typically forage for information by navigating from page to page along Web links. Their surfing patterns or actions are guided by their information needs. Researchers need tools to explore the complex interactions between user needs, user actions, and the structures and contents of the Web. In this paper, we describe two computational methods for understanding the relationship between user needs and user actions. First, for a particular pattern of surfing, we seek to infer ...

Keywords: World Wide Web, data mining, information foraging, information retrieval, information scent, usability

23 [Contributed articles: "In vivo" spam filtering: a challenge problem for KDD](#)

Tom Fawcett

 December 2003 [ACM SIGKDD Explorations Newsletter](#), Volume 5 Issue 2

Full text available:  pdf(260.66 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

Spam, also known as Unsolicited Commercial Email (UCE), is the bane of email communication. Many data mining researchers have addressed the problem of detecting spam, generally by treating it as a static text classification problem. True *in vivo* spam filtering has characteristics that make it a rich and challenging domain for data mining. Indeed, real-world datasets with these characteristics are typically difficult to acquire and to share. This paper demonstrates some of these characteri ...

Keywords: challenge problems, class skew, concept drift, cost-sensitive learning, data streams, imbalanced data, spam, text classification

24 AI update

September 2001 **intelligence**, Volume 12 Issue 3

Full text available:  pdf(129.12 KB)  html(46.75 KB) Additional Information: [full citation](#), [index terms](#)



25 An intelligent distributed environment for active learning

Yi Shang, Hongchi Shi, Su-Shing Chen

April 2001 **Proceedings of the tenth international conference on World Wide Web**

Full text available:  pdf(200.31 KB) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

Keywords: XML, active learning, multi-agent system, web-based education



26 Web mining for web personalization

Magdalini Eirinaki, Michalis Vazirgiannis

February 2003 **ACM Transactions on Internet Technology (TOIT)**, Volume 3 Issue 1

Full text available:  pdf(293.73 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Web personalization is the process of customizing a Web site to the needs of specific users, taking advantage of the knowledge acquired from the analysis of the user's navigational behavior (usage data) in correlation with other information collected in the Web context, namely, structure, content, and user profile data. Due to the explosive growth of the Web, the domain of Web personalization has gained great momentum both in the research and commercial areas. In this article we present a survey ...

Keywords: WWW, Web personalization, Web usage mining, user profiling



27 Gaussian process classification for segmenting and annotating sequences

Yasemin Altun, Thomas Hofmann, Alexander J. Smola

July 2004 **Twenty-first international conference on Machine learning**

Full text available:  pdf(204.35 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

Many real-world classification tasks involve the prediction of multiple, inter-dependent class labels. A prototypical case of this sort deals with prediction of a sequence of labels for a sequence of observations. Such problems arise naturally in the context of annotating and segmenting observation sequences. This paper generalizes Gaussian Process classification to predict multiple labels by taking dependencies between neighboring labels into account. Our approach is motivated by the desire to ...

28 Research track papers: Cyclic pattern kernels for predictive graph mining

Tamás Horváth, Thomas Gärtner, Stefan Wrobel

August 2004 **Proceedings of the 2004 ACM SIGKDD international conference on Knowledge discovery and data mining**Full text available:  [pdf\(291.65 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

With applications in biology, the world-wide web, and several other areas, mining of graph-structured objects has received significant interest recently. One of the major research directions in this field is concerned with predictive data mining in graph databases where each instance is represented by a graph. Some of the proposed approaches for this task rely on the excellent classification performance of support vector machines. To control the computational cost of these approaches, the underl ...

Keywords: computational chemistry, graph mining, kernel methods

29 User-cognizant multidimensional analysis

Sunita Sarawagi

September 2001 **The VLDB Journal — The International Journal on Very Large Data Bases**, Volume 10 Issue 2-3Full text available:  [pdf\(248.65 KB\)](#) Additional Information: [full citation](#), [abstract](#), [index terms](#)

Our goal is to enhance multidimensional database systems with a suite of advanced operators to automate data analysis tasks that are currently handled through manual exploration. In this paper, we present a key component of our system that characterizes the information content of a cell based on a user's prior familiarity with the cube and provides a context-sensitive exploration of the cube. There are three main modules of this component. A Tracker, that continuously tracks the parts of the cub ...

Keywords: Maximum entropy, Multidimensional data exploration, OLAP, Personalized mining, User-sensitive interest measure

30 Full papers: Iterative record linkage for cleaning and integration

Indrajit Bhattacharya, Lise Getoor

June 2004 **Proceedings of the 9th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery**Full text available:  [pdf\(264.99 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Record linkage, the problem of determining when two records refer to the same entity, has applications for both data cleaning (deduplication) and for integrating data from multiple sources. Traditional approaches use a similarity measure that compares tuples' attribute values; tuples with similarity scores above a certain threshold are declared to be matches. While this method can perform quite well in many domains, particularly domains where there is not a large amount of noise in the data, in ...

Keywords: clustering, deduplication, distance measure, record linkage

31 Strategic directions in electronic commerce and digital libraries: towards a digital agora

Nabil Adam, Yelena Yesha

December 1996 **ACM Computing Surveys (CSUR)**, Volume 28 Issue 4Full text available:  [pdf\(244.34 KB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

32 Scalable feature selection, classification and signature generation for organizing large text databases into hierarchical topic taxonomies

Soumen Chakrabarti, Byron Dom, Rakesh Agrawal, Prabhakar Raghavan

August 1998 **The VLDB Journal — The International Journal on Very Large Data Bases**,

Volume 7 Issue 3

Full text available:  [pdf\(281.37 KB\)](#) Additional Information: [full citation](#), [abstract](#), [citations](#), [index terms](#)

We explore how to organize large text databases hierarchically by topic to aid better searching, browsing and filtering. Many corpora, such as internet directories, digital libraries, and patent databases are manually organized into topic hierarchies, also called *taxonomies*. Similar to indices for relational data, taxonomies make search and access more efficient. However, the exponential growth in the volume of on-line textual information makes it nearly impossible to maintain such taxono ...

33 NSF workshop on industrial/academic cooperation in database systems

Mike Carey, Len Seligman

March 1999 **ACM SIGMOD Record**, Volume 28 Issue 1

Full text available:  [pdf\(1.96 MB\)](#) Additional Information: [full citation](#), [index terms](#)

34 Email classification with co-training

Svetlana Kiritchenko, Stan Matwin

November 2001 **Proceedings of the 2001 conference of the Centre for Advanced Studies on Collaborative research**

Full text available:  [pdf\(228.21 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#)

The main problems in text classification are lack of labeled data, as well as the cost of labeling the unlabeled data. We address these problems by exploring co-training - an algorithm that uses unlabeled data along with a few labeled examples to boost the performance of a classifier. We experiment with co-training on the email domain. Our results show that the performance of co-training depends on the learning algorithm it uses. In particular, Support Vector Machines significantly outperforms N ...

35 Development and use of a gold-standard data set for subjectivity classifications

Janyce M. Wiebe, Rebecca F. Bruce, Thomas P. O'Hara

June 1999 **Proceedings of the 37th conference on Association for Computational Linguistics**

Full text available:  [pdf\(744.73 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#)

This paper presents a case study of analyzing and improving intercoder reliability in discourse tagging using statistical techniques. Bias-corrected tags are formulated and successfully used to guide a revision of the coding manual and develop an automatic classifier.

36 Special issue on ICML: Coupled clustering: a method for detecting structural correspondence

Zvika Marx, Ido Dagan, Joachim M. Buhmann, Eli Shamir

March 2003 **The Journal of Machine Learning Research**, Volume 3

Full text available:  [pdf\(967.15 KB\)](#) Additional Information: [full citation](#), [abstract](#), [index terms](#)

This paper proposes a new paradigm and a computational framework for revealing equivalencies (analogies) between sub-structures of distinct composite systems that are

initially represented by unstructured data sets. For this purpose, we introduce and investigate a variant of traditional data clustering, termed *coupled clustering*, which outputs a configuration of corresponding subsets of two such representative sets. We apply our method to synthetic as well as textual data. Its achievement ...

37 Poster session 1: M/ORIS: a medical/operating room interaction system

Sébastien Grange, Terrence Fong, Charles Baur

October 2004 **Proceedings of the 6th international conference on Multimodal interfaces**

Full text available:  pdf(1.53 MB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

We propose an architecture for a real-time multimodal system, which provides non-contact, adaptive user interfacing for Computer-Assisted Surgery (CAS). The system, called M/ORIS (for Medical/Operating Room Interaction System) combines gesture interpretation as an explicit interaction modality with continuous, real-time monitoring of the surgical activity in order to automatically address the surgeon's needs. Such a system will help reduce a surgeon's workload and operation time. This paper f ...

Keywords: CAS, HCI, medical user interfaces, multimodal interaction

38 Learning classifiers: Using urls and table layout for web classification tasks

L. K. Shih, D. R. Karger

May 2004 **Proceedings of the 13th international conference on World Wide Web**

Full text available:  pdf(357.43 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

We propose new features and algorithms for automating Web-page classification tasks such as content recommendation and ad blocking. We show that the automated classification of Web pages can be much improved if, instead of looking at their textual content, we consider each link's URL and the visual placement of those links on a referring page. These features are unusual: rather than being scalar measurements like word counts they are *tree structured*---describing the position of the item ...

Keywords: classification, news recommendation, tree structures, web applications

39 An infrastructure for context-awareness based on first order logic

Anand Ranganathan, Roy H. Campbell

December 2003 **Personal and Ubiquitous Computing**, Volume 7 Issue 6

Full text available:  pdf(319.19 KB) Additional Information: [full citation](#), [abstract](#), [index terms](#)

Context simplifies and enriches human-human interaction. However, enhancing human-computer interaction through the use of context remains a difficult task. Applications in pervasive and mobile environments need to be context-aware so that they can adapt themselves to rapidly changing situations. One of the problems is that there is no common, reusable model for context used by these environments. In this paper, we propose a model of context that is based on first order predicate calculus. The fi ...

Keywords: Context-awareness, Infrastructure, Logic

40 Image Categorization by Learning and Reasoning with Regions

Yixin Chen, James Z. Wang

August 2004 **The Journal of Machine Learning Research**, Volume 5

Full text available:  pdf(1.31 MB) Additional Information: [full citation](#), [abstract](#)

Designing computer programs to automatically categorize images using low-level features is

a challenging research topic in computer vision. In this paper, we present a new learning technique, which extends Multiple-Instance Learning (MIL), and its application to the problem of region-based image categorization. Images are viewed as bags, each of which contains a number of instances corresponding to regions obtained from image segmentation. The standard MIL problem assumes that a bag is labeled p ...

Results 21 - 40 of 200

Result page: [previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)

Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

by strong research interests as well as urgent practical, social, and economical needs. While the last few years knowledge discovery tools have been used mainly in research environments, sophisticated software products are now rapidly emerging. In this paper, we provide an overview of common knowledge discovery tasks and approaches to solve these tasks. We propose a feature classification scheme that can be ...

Keywords: data mining, knowledge discovery in databases, surveys

44 Constraints in data mining: SPARTAN: using constrained models for guaranteed-error semantic compression 

Shivnath Babu, Minos Garofalakis, Rajeev Rastogi

June 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 1

Full text available:  [pdf\(259.12 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

While a variety of lossy compression schemes have been developed for certain forms of digital data (e.g., images, audio, video), the area of lossy compression techniques for arbitrary data tables has been left relatively unexplored. Nevertheless, such techniques are clearly motivated by the ever-increasing data collection rates of modern enterprises and the need for effective, guaranteed-quality approximate answers to queries over massive relational data sets. In this paper, we propose *SPARTAN* ...

45 A survey on wavelet applications in data mining 

Tao Li, Qi Li, Shenghuo Zhu, Mitsunori Ogihara

December 2002 **ACM SIGKDD Explorations Newsletter**, Volume 4 Issue 2

Full text available:  [pdf\(330.06 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

Recently there has been significant development in the use of wavelet methods in various data mining processes. However, there has been written no comprehensive survey available on the topic. The goal of this paper is to fill the void. First, the paper presents a high-level data-mining framework that reduces the overall process into smaller components. Then applications of wavelets for each component are reviewed. The paper concludes by discussing the impact of wavelets on data mining research and ...

46 Statistical methods I: Bayesian analysis of massive datasets via particle filters 

Greg Ridgeway, David Madigan

July 2002 **Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(896.64 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Markov Chain Monte Carlo (MCMC) techniques revolutionized statistical practice in the 1990s by providing an essential toolkit for making the rigor and flexibility of Bayesian analysis computationally practical. At the same time the increasing prevalence of massive datasets and the expansion of the field of data mining has created the need to produce statistically sound methods that scale to these large problems. Except for the most trivial examples, current MCMC methods require a complete scan or ...

47 Industrial/government track: Empirical Bayesian data mining for discovering patterns in post-marketing drug safety 

David M. Fram, June S. Almenoff, William DuMouchel

August 2003 **Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(461.25 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Because of practical limits in characterizing the safety profiles of therapeutic products prior to marketing, manufacturers and regulatory agencies perform post-marketing surveillance

based on the collection of adverse reaction reports ("pharmacovigilance"). The resulting databases, while rich in real-world information, are notoriously difficult to analyze using traditional techniques. Each report may involve multiple medicines, symptoms, and demographic factors, and there is no easily linked inf ...

Keywords: association rules, data mining, empirical Bayes methods, pharmacovigilance, post-marketing surveillance

48 Long papers: smart environments and ubiquitous computing: CASIS: a context-aware speech interface system 

Lee Hoi Leong, Shinsuke Kobayashi, Noboru Koshizuka, Ken Sakamura

January 2005 **Proceedings of the 10th international conference on Intelligent user interfaces**

Full text available:  [pdf\(530.98 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

In this paper, we propose a robust natural language interface called CASIS for controlling devices in an intelligent environment. CASIS is novel in a sense that it integrates physical context acquired from the sensors embedded in the environment with traditionally used context to reduce the system error rate and disambiguate deictic references and elliptical inputs. The n-best result of the speech recognizer is re-ranked by a score calculated using a Bayesian network consisting of information fr ...

Keywords: Bayesian network, context-aware computing, natural language processing, speech user interface

49 Automated assistants to aid humans in understanding team behaviors 

Taylor Raines, Milind Tambe, Stacy Marsella

June 2000 **Proceedings of the fourth international conference on Autonomous agents**

Full text available:  [pdf\(1.09 MB\)](#) Additional Information: [full citation](#), [references](#), [citations](#), [index terms](#)

50 Fast detection of communication patterns in distributed executions 

Thomas Kunz, Michiel F. H. Seuren

November 1997 **Proceedings of the 1997 conference of the Centre for Advanced Studies on Collaborative research**

Full text available:  [pdf\(4.21 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Understanding distributed applications is a tedious and difficult task. Visualizations based on process-time diagrams are often used to obtain a better understanding of the execution of the application. The visualization tool we use is Poet, an event tracer developed at the University of Waterloo. However, these diagrams are often very complex and do not provide the user with the desired overview of the application. In our experience, such tools display repeated occurrences of non-trivial commun ...

51 Data mining: an experimental undergraduate course 

Youmin Lu, Jennifer Bettine

February 2003 **Journal of Computing Sciences in Colleges**, Volume 18 Issue 3

Full text available:  [pdf\(28.15 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

Data mining is the extraction of implicit, previously unknown, and potentially useful information from data. Advances in information technology and data collection methods have led to the availability of large data sets in commercial enterprises and in a wide variety of scientific and engineering disciplines. We have an unprecedented opportunity to

analyze this data and extract intelligent and useful information. Traditionally, machine learning is a part of the Artificial Intelligence course. Up ...

52 Unsupervised Bayesian visualization of high-dimensional data

Petri Kontkanen, Jussi Lahtinen, Petri Myllymäki, Henry Tirri

August 2000 **Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(160.91 KB\)](#) Additional Information: [full citation](#), [references](#), [index terms](#)



53 Technique for automatically correcting words in text

Karen Kukich

December 1992 **ACM Computing Surveys (CSUR)**, Volume 24 Issue 4

Full text available:  [pdf\(6.23 MB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index terms](#), [review](#)

Research aimed at correcting words in text has focused on three progressively more difficult problems: (1) nonword error detection; (2) isolated-word error correction; and (3) context-dependent word correction. In response to the first problem, efficient pattern-matching and n-gram analysis techniques have been developed for detecting strings that do not appear in a given word list. In response to the second problem, a variety of general and application-specific spelling cor ...

Keywords: n-gram analysis, Optical Character Recognition (OCR), context-dependent spelling correction, grammar checking, natural-language-processing models, neural net classifiers, spell checking, spelling error detection, spelling error patterns, statistical-language models, word recognition and correction



54 Decomposable modeling in natural language processing

Rebecca F. Bruce, Janyce M. Wiebe

June 1999 **Computational Linguistics**, Volume 25 Issue 2

Full text available:  [pdf\(921.88 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)
 [Publisher Site](#)



In this paper, we describe a framework for developing probabilistic classifiers in natural language processing. Our focus is on formulating models that capture the most important interdependencies among features, to avoid overfitting the data while also characterizing the data well. The class of probability models and the associated inference techniques described here were developed in mathematical statistics, and are widely used in artificial intelligence and applied statistics. Our goal is to ...

55 Towards automated synthesis of data mining programs

Wray Buntine, Bernd Fischer, Thomas Pressburger

August 1999 **Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining**

Full text available:  [pdf\(637.67 KB\)](#) Additional Information: [full citation](#), [references](#), [index terms](#)



56 Tree induction vs. logistic regression: a learning-curve analysis

Claudia Perlich, Foster Provost, Jeffrey S. Simonoff

December 2003 **The Journal of Machine Learning Research**, Volume 4

Full text available:  [pdf\(263.37 KB\)](#) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#), [index](#)



terms

Tree induction and logistic regression are two standard, off-the-shelf methods for building models for classification. We present a large-scale experimental comparison of logistic regression and tree induction, assessing classification accuracy and the quality of rankings based on class-membership probabilities. We use a learning-curve analysis to examine the relationship of these measures to the size of the training set. The results of the study show several things. (1) Contrary to some prior o ...

57 Position papers on MRDM: Prospects and challenges for multi-relational data mining 

Pedro Domingos

July 2003 **ACM SIGKDD Explorations Newsletter**, Volume 5 Issue 1Full text available:  pdf(397.89 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [citations](#)

This short paper argues that multi-relational data mining has a key role to play in the growth of KDD, and briefly surveys some of the main drivers, research problems, and opportunities in this emerging field.

58 Position papers: Theoretical frameworks for data mining 

Heikki Mannila

January 2000 **ACM SIGKDD Explorations Newsletter**, Volume 1 Issue 2Full text available:  pdf(341.62 KB) Additional Information: [full citation](#), [references](#), [citations](#)**59 Learning Bayesian network classifiers by maximizing conditional likelihood** 

Daniel Grossman, Pedro Domingos

July 2004 **Twenty-first international conference on Machine learning**Full text available:  pdf(187.23 KB) Additional Information: [full citation](#), [abstract](#), [references](#)

Bayesian networks are a powerful probabilistic representation, and their use for classification has received considerable attention. However, they tend to perform poorly when learned in the standard way. This is attributable to a mismatch between the objective function used (likelihood or a function thereof) and the goal of classification (maximizing accuracy or conditional likelihood). Unfortunately, the computational cost of optimizing structure and parameters for conditional likelihood is pro ...

60 Research papers: data mining: Pre-empting user questions through anticipation: data mining FAQ lists 

Dick Ng'Ambi

September 2002 **Proceedings of the 2002 annual research conference of the South African institute of computer scientists and information technologists on Enablement through technology**Full text available:  pdf(202.21 KB) Additional Information: [full citation](#), [abstract](#), [references](#), [index terms](#)

In this paper we describe the use of data mining techniques on frequently referenced questions (FRQ) to predict the user's 'next' question with the view to pre-empting the question using proactive response. Relationships and patterns hidden in frequently asked questions (FAQ) lists, once discovered, can be used to anticipate user questions and enrich the questioning engagement. A prototype, dynamic Intelligent Handler of Frequently Asked Questions, has been developed to help predict user questio ...

Keywords: associative rule, data mining, dynamic FAQ lists, intelligent frequently asked questions, pre-empting

Results 41 - 60 of 200

Result page: [previous](#) [1](#) [2](#) **3** [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [next](#)

The ACM Portal is published by the Association for Computing Machinery. Copyright © 2005 ACM, Inc.

[Terms of Usage](#) [Privacy Policy](#) [Code of Ethics](#) [Contact Us](#)Useful downloads:  [Adobe Acrobat](#)  [QuickTime](#)  [Windows Media Player](#)  [Real Player](#)

USPTO PATENT FULL-TEXT AND IMAGE DATABASE

[Home](#)[Quick](#)[Advanced](#)[Pat Num](#)[Help](#)[Bottom](#)[View Cart](#)

Searching 1976 to present...

Results of Search in 1976 to present db for:

((("natural language" AND "data mining") AND goal) AND Bayesian): 6 patents.

Hits 1 through 6 out of 6

[Jump To](#)[Refine Search](#)

"natural language" and "data mining" and goal and Bay

PAT. NO. Title

1 [6,850,252](#) [T Intelligent electronic appliance system and method](#)

2 [6,820,075](#) [T Document-centric system with auto-completion](#)

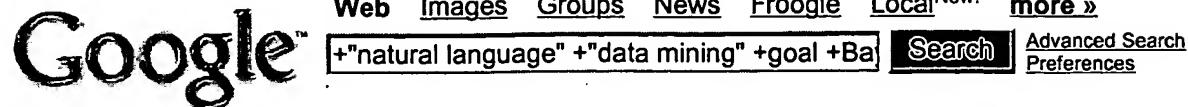
3 [6,778,979](#) [T System for automatically generating queries](#)

4 [6,732,090](#) [T Meta-document management system with user definable personalities](#)

5 [6,611,825](#) [T Method and system for text mining using multidimensional subspaces](#)

6 [6,233,575](#) [T Multilevel taxonomy based on features derived from training documents classification using fisher values as discrimination values](#)

[Top](#)[View Cart](#)[Home](#)[Quick](#)[Advanced](#)[Pat Num](#)[Help](#)

**Web**Results 1 - 10 of about 7,660 for **+"natural language" +"data mining" +goal +Bayesian.** (0.35 seconds)**Artificial Intelligence Lab**

The ultimate **goal** of GeneScene is to provide with two different approaches of **natural language** processing (NLP The **data mining** component is designed to extract ...
ai.bpa.arizona.edu/research/bioinformatics/ - 52k - Cached - Similar pages

Sponsored Links

Bayesian Data Mining
 a powerful technology for extracting useful knowledge
www.bayesia.com

Data Mining and Discovery

'By contrast, the **goal** of **data** group "is focused on learning from data and **data mining**. ... such as handwriting recognition and **natural-language** processing, and (2 ...
www.aaai.org/AITopics/html/mining.html - 44k - Cached - Similar pages

[1-13] Glossary of AI terms.

... distance remaining from the current node to the **goal**. Also called **natural language** processing. **Data Mining**: Also known as Knowledge Discovery in Databases (KDD ...
www.faqs.org/faqs/ai-faq/general/part1/section-14.html - 12k - Cached - Similar pages

Kevin B. Korb

... in Knowledge Discovery and **Data Mining** (1998 New Directions in Planning and **Natural Language** Generation, ECCAI AAAI spring symposium series: **Goal**-driven learning ...
www.csse.monash.edu.au/~korb/ - 22k - Feb 5, 2005 - Cached - Similar pages

[PPT] Modeling a Concept-Based Information System to Promote Scientific ...

File Format: Microsoft Powerpoint 97 - View as HTML
Goal: links between the findings documented in the literature and the GENIES: a **natural language** processing system for the extraction of molecular **Data Mining**.
www.telemakus.net/papers/BioarrayMtg_11-20-03.ppt - Similar pages

Daniel L. Rubin, MD, MS

... for feature extraction and **data mining** in this The **goal** is to integrate anatomic knowledge ... interested in developing statistical **natural language** processing (NLP ...
smi-web.stanford.edu/people/rubin/DLRresearch.html - 10k - Cached - Similar pages

Center for Automated Learning and Discovery

The **goal** of her research is to seek machine learning such as understanding **natural language** text, and ... works on the application of **data mining** and statistical ...
www.cald.cs.cmu.edu/masters/faculty.page.html - 63k - Cached - Similar pages

[PPT] Mining Data Discussion Data in Virtual Learning Communities

File Format: Microsoft Powerpoint 97 - View as HTML
 ... intelligence and computational sciences fall short in their ability to support **natural language** understanding. **Data Mining** Process. ... Determine research **goal**.
www.icgd.usask.ca/DanielGradWorkshop2004.ppt - Similar pages

ICAIL 2005 Workshop on Data Mining, Information Extraction, and ...

... of data analysis include **data mining** (discovering relevant ... AI) techniques, such as **natural language** understanding, **Bayesian** ... The **goal** of this workshop will be ...
www.karlbranting.net/icail2005workshopcfp.html - 7k - Cached - Similar pages

ECML/PKDD-2002 Tutorial: Text Mining and Internet Content ...

This **goal** will be achieved by realizing the concepts about the field through The role of

lexicons in **natural language** processing. ... Untangling text **data mining**.
www.cs.helsinki.fi/events/ecmlpkdd/hidalgo.html - 19k - Feb 6, 2005 -
[Cached](#) - [Similar pages](#)

Gooooooooogle ►

Result Page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [Next](#)

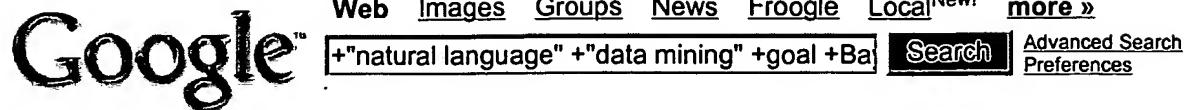


Free! [Google Desktop Search](#): Search your own computer.

[Search within results](#) | [Language Tools](#) | [Search Tips](#) | [Dissatisfied? Help us improve](#)

[Google Home](#) - [Advertising Programs](#) - [Business Solutions](#) - [About Google](#)

©2005 Google

**Web**

Results 11 - 20 of about 7,660 for +"natural language" +"data mining" +goal +Bayesian. (0.30 seconds)

Abstracts

"My main research **goal** is to interests include machine learning and **data mining** (of web ... liked by others), statistical **natural language** processing, multiple ... learning.cis.upenn.edu/ - 14k - Cached - Similar pages

Sponsored Links

Bayesian Data Mining
a powerful technology for extracting useful knowledge
www.bayesia.com

[PDF] What is Data MiningFile Format: PDF/Adobe Acrobat - View as HTML

Page 8. Venkat Chalasani SRA **Data Mining** Tasks **Mortgage Loan Goal** : To grant or reject loan application Support vector machines • **Bayesian** networks ... www.altum.com/bcig/events/tutorials/2003_05.pdf - Similar pages

[PPT] Bayesian Networks and the Semantic WebFile Format: Microsoft Powerpoint 97 - View as HTML

diagnostic aids, classification., **Natural language** desambiguation., **data mining**. Ontologies mapping ? **Goal**. Convert an OWL ontology to a **Bayesian** Network. www.deri.at/teaching/seminars/internal/slides/Bayesian%20Networks%20and%20the%20Semantic%20Web.ppt - Similar pages

[PDF] Data Mining Analytics for Business Intelligence and Decision ...File Format: PDF/Adobe Acrobat - View as HTML

... handwriting recognition, and **natural language** understanding privacy preserving **data mining**, and heterogeneous **data mining**. The key **goal** in solutions research is ... www.research.ibm.com/dar/papers/pdf/orms2.pdf - Similar pages

AI, CogSci and Robotics: Artificial Intelligence

... in agents with **goal** autonomy, coordinating applications such as **natural language** acquisition, problem MLE), machine learning, **data mining**, pattern recognition ... transit-port.net/AI.CogSci.Robotics/ai.html - 50k - Feb 6, 2005 - Cached - Similar pages

[PPT] 基因體計畫 – 2002 yearsFile Format: Microsoft Powerpoint 97 - View as HTML

Future Work in Biological Knowledge Base and Agent (I). **Goal**: To facilitate scientific discovery. Genome rearrangements. **Natural language** processing. ... **Data mining**. 140.109.19.57/tigbp/04-05Fall/orientation/sung%20.ppt - Similar pages

The Use of Java in Machine Learning

... they will be able to understand (**Natural Language** Representation The **goal** of docking to the battery charger ILP is used successfully in **Data Mining** for finding ... www.developer.com/java/other/article.php/10936_1559871_2 - 54k - Cached - Similar pages

Sato Lab, Tokyo Institute of Technology

... for parameterized logic programs by inter-goal sharing. Journal of **Natural Language** Processing, Vol.8, No.1 are quite powerful tools for **data mining**, but there ... sato-www.cs.titech.ac.jp/en/research/ - 13k - Cached - Similar pages

Surfmind.com Web Collection: Cognitive Science Meets the Web

... on Knowledge Discovery and **Data Mining**, Newport Beach, CA. Maps of words for **natural language** processing applications Emergence of a Design **Goal** and Theoretical ...

www.surfmind.com/web/searchresult.cfm?criteria=classification+and+artificial+intelligence -
62k - [Cached](#) - [Similar pages](#)

Tutorial AI*IA'99

... to spatial data analysis (**Spatial Data Mining**), where spatial di Roma Tor Vergata,
Natural Language Processing Group The **goal** of the tutorial will be to present ...
www.di.uniba.it/~malerba/aiia99/tutorial_eng.html - 19k - [Cached](#) - [Similar pages](#)

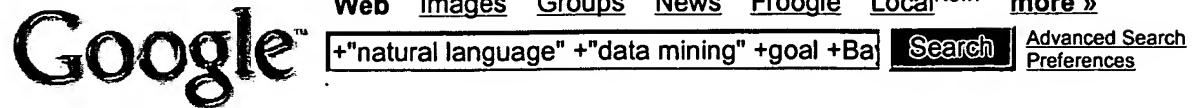
◀ Gooooooooooooogle ▶

Result Page: [Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [Next](#)

[Search within results](#) | [Language Tools](#) | [Search Tips](#)

[Google Home](#) - [Advertising Programs](#) - [Business Solutions](#) - [About Google](#)

©2005 Google

**Web**

Results 21 - 30 of about 7,660 for +"natural language" +"data mining" +goal +Bayesian. (0.29 seconds)

[PDF] From Data Mining to Knowledge Discovery in DatabasesFile Format: PDF/Adobe Acrobat - [View as HTML](#)

... is useful for understanding the overall discovery **goal**. The relative importance of prediction and description for particular **data-mining** applications can ...
www.kdnuggets.com/gspubs/aimag-kdd-overview-1996-Fayyad.pdf - [Similar pages](#)

Sponsored Links

Bayesian Data Mining
 a powerful technology for extracting useful knowledge
www.bayesia.com

Applied and Interdisciplinary Informatics Group

... Tutoring Systems, Expert Systems, **natural language** processing/language Nicholas Miles, Applying **data mining** using evolutionary ... order to achieve this **goal** I am ...
www.cs.kent.ac.uk/research/groups/aiii/membership.html - 22k - Feb 5, 2005 -
[Cached](#) - [Similar pages](#)

IEEE Transactions on Knowledge and Data Engineering, Volume 14

... for Understanding Domain-Specific **Natural Language** Queries A Method for Clustering Objects for Spatial **Data Mining**. Probability Bounds for **Goal** Directed Queries ...
www.informatik.uni-trier.de/~ley/db/journals/tkde/tkde14.html - 59k - Feb 4, 2005 -
[Cached](#) - [Similar pages](#)

NEuroNet: Roadmap: Selected Applications - Intelligent Agents

Issues regarding speech recognition and **natural language** understanding are indexing (LSI) method towards the same **goal**. in the field of **data mining** in general.
www.kcl.ac.uk/neuronet/about/roadmap/intelagts.html - 30k - Feb 5, 2005 -
[Cached](#) - [Similar pages](#)

Research Topics

The central **goal** of the project is to develop new to devise privacy-preserving methods for **data mining** from sources **Natural Language Processing** and **Text Mining**.
www.di.unipi.it/ricerca/Report2003/node199.html - 17k - [Cached](#) - [Similar pages](#)

[PDF] Bayesian Networks and the Semantic WebFile Format: PDF/Adobe Acrobat - [View as HTML](#)

diagnostic aids, classification, **Natural language** desambiguation, **data mining** Ontologies mapping ? information **Goal** • Convert an OWL ontology to a **Bayesian** ...
www.deri.at/teaching/seminars/internal/slides/Bayesian%20Networks%20and%20the%20Semantic%20Web.pdf - [Similar pages](#)

Research Subareas

It will also be a **goal** of this area to, In **natural language** dialogues system we expect to build a module as well as automated learning and **data mining**, so as to ...
centria.di.fct.unl.pt/subareas.html - 12k - [Cached](#) - [Similar pages](#)

[PDF] Natural Language Processing and Systems BiologyFile Format: PDF/Adobe Acrobat - [View as HTML](#)

Examples of papers whose stated **goal** is database **Natural Language Processing** and **Systems Biology** 9 have included decision trees, **Bayesian** classifiers, hidden ...
compbio.uchsc.edu/Hunter_lab/Cohen/Cohen.pdf - [Similar pages](#)

[PDF] Ganesh RamakrishnanFile Format: PDF/Adobe Acrobat - [View as HTML](#)

The **goal** was to identify images of new NIS Journal, Special Issue on **Natural Language Interface** to to The SIAM International Conference on **Data Mining**, 2005 ...
www.cse.iitb.ac.in/~hare/reference/ganesh-resume.pdf - [Similar pages](#)

Finding Frequent Substructures in Chemical Compounds

... rather than based on classification; the **goal** being to that fits in the **descriptive data mining** paradigm We apply WARMR to the **natural language processing** task ...
www.cs.kuleuven.ac.be/~ldh/publicaties/abstracts.html - 13k - [Cached](#) - [Similar pages](#)

◀ Goooooooooooooogle ▶

Result Page: [Previous](#) [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [11](#) [12](#) [Next](#)

[Search within results](#) | [Language Tools](#) | [Search Tips](#)

[Google Home](#) - [Advertising Programs](#) - [Business Solutions](#) - [About Google](#)

©2005 Google